CPP 104: Topical Session: Data Driven Materials Science - Machine Learning Applications (joint session MM/CPP)

Time: Thursday 17:30–19:00   Location: BAR 205

CPP 104.1 Thu 17:30 BAR 205

How polymorphism of adsorbate molecules determines the physical properties of metal/organic interfaces: a large scale study — •Johannes J. Cartus, Andreas Jeindl, Lukas Hörmann, and Oliver T. Hofmann — Institute of Solid State Physics, NAWI Graz, Graz University of Technology, Graz, Austria

The work function of metallic surfaces can be readily tuned via the interface dipole that emerges when adsorbing organic electron-donor or -acceptor molecules. Despite extensive research in the field and a wealth of experimental data available, it is presently still not possible to predict the magnitude of the dipole based on the choice of the interface materials alone. This is because molecular properties and the structure of adsorbate layers are rarely considered simultaneously, even though both are known to have great impact.

In this work, we present a systematic study of selected substrate/adsorbate combinations, for which we computationally determine the electronic structure of various surface polymorphs. This allows us to explore which kind of structures preferentially form for different molecule classes and directly relate them to the interface work function and other system properties.

CPP 104.2 Thu 17:45 BAR 205

Investigation of short-range order in multicomponent alloys with the use of machine-learning interatomic potentials — •Tatiana Kostiuchenko[1], Alexander Shapeev[1], Fritz Körmann[2,3], and Andrey Ruban[4,5] — [1]Skolkovo Institute of Science and Technology, Moscow, Russia — [2]Delft University of Technology, Delft, The Netherlands — [3]Max-Planck-Institut für Eisenforschung GmbH, Düsseldorf, Germany — [4]KTH Royal Institute of Technology, Stockholm, Sweden — [5]Materials Center Leoben Forschung GmbH, Leoben, Austria

Multicomponent alloys, such as high-entropy alloys or multi-principal element alloys, are promising structural materials. However, the vast range of chemical compositions and long annealing times offer an opportunity for their study with ab initio methods. The drawback of ab initio methods is their huge computational costs. We address this drawback by developing efficient data-driven interaction models with accuracy close to those of ab initio methods, namely the low-rank interatomic potential (LRP) [Shapeev A., 2017]. They are used in a Canonical Monte Carlo algorithm as an "on-lattice" model that can take into account local lattice distortions. In this work, we investigate the equiatomic VCoNi system. It represents a medium-entropy alloy with a distorted fcc lattice which leads to an outstanding strength-ductility relationship as reported in [Sohn S. et al., 2019]. We simulate this system by including, implicitly, magnetism into LRP by fitting on spin-polarized DFT calculation. The LRP has the error of about 1-3 meV/atom and is used to study the formation of the short-range order.

CPP 104.3 Thu 18:00 BAR 205

An equation for membrane permeability: Insight from compressed sensing — •Arghya Dutta and Tristan Bereau — Max Planck Institute for Polymer Research, Mainz, Germany

Using a material's structure and readily available properties to predict a difficult-to-measure but important property is crucial in natural sciences and engineering. Data mining—the process of discovering associations, correlations, and anomalies in data—can significantly facilitate the search for these generalized structure-property relationships by providing relevant descriptors. To give an example, the efficacy of a targeted drug depends on whether or not it can go through a cell membrane. This capacity is quantified by permeability which measures the drug's flux across the membrane. However, calculating permeability is computationally expensive. In this presentation, I will discuss results from our ongoing search for a simple and interpretable equation, which is a function of key physical descriptors, for permeability using compressed sensing methods. This simplified description of permeability will allow simulation-free prediction, and, potentially, assist in rapid screening of candidate drug molecules.

CPP 104.4 Thu 18:15 BAR 205

Transferable Gaussian Process Regression for prediction of

molecular crystals harmonic free energy. — •Marcin Krynski[1] and Mariana Rossi[1,2] — [1]Fritz Haber Institute of the Max Planck Society, Berlin, Germany — [2]MPI for Structure and Dynamics of Matter, Hamburg, Germany

Organic molecular crystals are a large group of compounds with properties tied strongly to the crystallographic structure of their numerous polymorphs. While thermodynamic free energies are necessary for obtaining a reliable polymorph energy ranking[1,2], their inclusion in large-scale simulations for polymorph screening is challenging, because dispersion-corrected DFT accuracy is needed in order to capture the complex charge rearrangement and bond-softening. Therefore, to predict harmonic Helmholtz free energies, we devised a framework that employs the transferable Gaussian Process Regression model with Smooth Overlap of Atomic Positions[3] descriptors representing local atomic environments. We developed strategies based on farthest point sampling to minimize the size of the training set and to ensure statistical diversity. We benchmark our framework on a set of 444 hydrocarbon crystal polymorphs. Superior performance and high prediction accuracy, with mean absolute deviation below 0.1 meV/atom is achieved by a hyperparameter optimisation performed on empirical-potential models, ensuring sensitivity to longer-range structural patterns.

[1] J. Nyman and G. Day, *CrystEngComm* **17**, 5154 (2015);
[2] M. Rossi, P. Gasparotto, M. Ceriotti, *PRL* **117**, 115702 (2016);
[3] A. P. Bartók, R. Kondor, G. Csányi, *PRB* **87**, 184115 (2013).

CPP 104.5 Thu 18:30 BAR 205

Bayesian modeling for potential energy surface minimization — •Estefania Garijo del Rio, Sami Juhani Kaapa, and Karsten Wedel Jacobsen — CAMD, Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

Computational simulations using electronic structure methods of materials and molecules require the (meta-)stable structure of the system under investigation to be known. In the absence of experimental structural data, the usual procedure is to use quantum chemistry codes together with some optimization algorithm to find successive approximations of a (local) minimum of the potential energy surface under the Born-Oppenheimer approximation. In this context, methods that incorporate machine learning surrogate models that are built on the fly to reduce the number of evaluations have recently gained popularity. Here, we explore and compare how different choices for the kernels can affect the performance of the optimization when Gaussian process regression is used to fit energies and forces.

CPP 104.6 Thu 18:45 BAR 205

A computational route between band mapping and band structure — R. Patrick Xian[1], •Vincent Stimper[2], Marios Zacharias[1], Shuo Dong[1], Maciej Dendzik[1], Samuel Beaulieu[1], Matthias Scheffler[1], Bernhard Schölkopf[2], Martin Wolf[1], Laurenz Rettig[1], Christian Carbogno[1], Stefan Bauer[2], and Ralph Ernstorfer[1] — [1]Fritz Haber Institute of the Max Planck Society, Berlin, Germany — [2]Max Planck Institute for Intelligent Systems, Tübingen, Germany

In solid state physics, the electronic band structure imprints the multidimensional relations between energy and momentum of periodically confined electrons. Photoemission spectroscopy has provided a reliable source of experimental validation for electronic structure theory and the understanding of electronic properties. Recent advancements in UV sources and detector technology leads to the growth of band mapping data in resolution, size and scale. This offers fresh opportunities for high-throughput material characterization. Here, we present an efficient probabilistic machine learning framework to connect photoemission band mapping with band structure calculations by reconstructing the full 3D valence band structure of tungsten diselenide within its first Brillouin zone. We then digitize the empirical shape of the reconstructed bands using orthonormal bases for multiscale sampling and comparison with theory. Our approaches open up avenues for the precise quantification of band structures and realize a viable path integrating band mapping data with computational materials science databases.